

Amino-terminal amino acid sequence of human leukocyte interferon

(micro sequence determination/interferon diversity/automatic sequence determination/multigene family/
protein primary structure)

WARREN P. LEVY*, JOHN SHIVELY†, MENACHEM RUBINSTEIN*‡, URSINO DEL VALLE†, AND
SIDNEY PESTKA*

*Roche Institute of Molecular Biology, Nutley, New Jersey 07110; and †Division of Immunology, City of Hope Research Institute, Duarte, California 91010

Communicated by Sidney Udenfriend, June 2, 1980

ABSTRACT We report the amino-terminal sequence of the first 22 amino acids of human leukocyte interferon. These and other results indicate that human leukocyte interferon consists of many individual species. We, therefore, postulate that diversity in this protein is routinely present and that the human leukocyte interferons represent a multigene family.

Although 23 years have passed since the discovery of interferon (1), purification of several interferons has been attained only in the last few years. We previously reported the purification and amino acid composition of human leukocyte interferon (2). Within the past few months, the amino-terminal amino acid sequences of several interferons were reported (3-5). The major obstacle in determining the sequence of interferon has been the availability of sufficient amounts of pure material. Recent advances in manual (6, 7) and automatic (8, 9) sequence analysis have made possible the determination of the primary structure of microgram amounts of protein. We report here the amino-terminal sequence of human leukocyte interferon as determined by automatic and manual sequencing of the native protein and tryptic fragments. In addition, we extend the recently reported amino-terminal sequence and describe two significant differences between our data and those of Zoon *et al.* (4).

EXPERIMENTAL PROCEDURES

The purification of human leukocyte interferon from chronic myelogenous leukemia (CML) cells has been described (2, 10, 11). The sequences of human leukocyte interferon species α_1 and β_2 were determined with a modified Beckman spinning-cup sequenator. Peptide fragments of α_1 , α_2 , and β_1 species were produced by digestion with trypsin and purified by high-performance liquid chromatography (HPLC). The sequences of tryptic peptides were determined manually by the Edman reaction (12), and the phenylthiohydantoin derivatives of the amino acids were identified by HPLC (13-15).

Automatic Edman degradations were performed in a modified Beckman 890C sequenator on 1.7 nmol of species α_1 of human leukocyte interferon. The modifications, which are similar to those described by Wittmann-Liebold (8) and Hunziker and Hood (16), include an improved vacuum system, improved reagent and solvent delivery system, extensive solvent and reagent purification, and a device (17) that automatically converts anilinothiazolinone to phenylthiohydantoin derivatives of amino acids. Proteins are retained in the spinning cup with 6 mg of Polybrene, which, together with 100 nmol of glycyl-

glycine, has been subjected to seven precycles of Edman degradation. Phenylthiohydantoin amino acids were analyzed by HPLC on Du Pont Zorbax octadecylsilica or cyanopropylsilica columns on a Waters Associates chromatograph, by monitoring absorbance at 254 nm and 313 nm. Peak assignments, except for serine, were made by chromatography on a Zorbax octadecylsilica column. Phenylthiohydantoin serine was identified as the "dehydro" derivative on a cyanopropylsilica column. Peaks were integrated and gradient elution was controlled by a Spectra Physics SP4000 integration system. All phenylthiohydantoin derivatives were detected by their absorbance at 254 nm, except for those of serine and threonine, which were detected at 313 nm.

RESULTS AND DISCUSSION

The compositions of the first two tryptic peptides of species α_1 , α_2 , and β_1 of human leukocyte interferon are shown in Table 1. The amino-terminal peptide had been identified before sequence data were available by analysis of a tryptic digest of interferon (α_2) that was blocked at the amino terminus and, therefore, would not react with fluorescamine. The sequence of peptide 2 was determined manually, and this peptide was identified as the penultimate tryptic peptide by comparison with the sequence determined by the automatic sequenator (Fig. 1). Manual Edman degradation provided the sequence of the first four amino acids as well as the terminal arginine (Table 1).

Fig. 1 shows the amino-terminal sequence analysis of 1.7 nmol of human leukocyte interferon species α_1 . On the basis of the yield of leucine at cycle 3 (540 pmol) and a repetitive yield of 92% (calculated from the leucines at cycles 3 and 9), the yield of the amino-terminal serine was calculated to be 41%. The high yield of the amino terminus and the determination of a single sequence from residues 1 through 21 confirms that this protein is a pure single species. The assignments of Asn-11, Arg-12, Thr-14, and Leu-18 were made with less confidence, because their yields were somewhat lower than expected. However, it should be noted that no other phenylthiohydantoin amino acid was seen for these cycles. Although our identification of Thr-14 is tentative from the data by automatic sequencing (Fig. 1), it is clear that alanine [reported for this position in lymphoblastoid interferon (4)] is not found at this position in human leukocyte interferon species α_1 . It is evident from the manually determined sequence (Table 1) that the amino-terminal threonine of peptide 2 must correspond to position 14. Peptide 2 was the only one whose composition and

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

Abbreviation: HPLC, high-performance liquid chromatography.

‡ Present address: The Weizmann Institute of Science, Rehovot, Israel.

Table 1. Compositions of the first two amino-terminal tryptic peptides and partial sequence of the second

Peptide	Composition
1	2 Asx, Thr, 2 Ser, Glx, Gly, 2 Leu, His, 2 Arg, Pro
2	Thr, Glx, Ala, 2 Met, 3 Leu, Arg

Sequence	
2	Thr-Leu-Met-Leu(Leu, Glx, Ala, Met)Arg

Compositions were determined by amino acid analysis with a fluorescamine analyzer. The sequence of the first four amino acids and the carboxy-terminal arginine of 1.4 nmol of peptide 2 (from α_2) were determined by manual Edman degradation. The first four were identified as phenylthiohydantoin derivatives. The terminal arginine was identified by amino acid analysis (after hydrolysis) of the remaining components after cycle 8. The sequence of the amino acids given in parentheses was not determined by the manual reaction. Peptide 1 was isolated from the tryptic digest of leukocyte interferon species α_2 .

sequence matched positions 14-22. Furthermore, although Arg-22 was not identified on the automatic sequencer (Fig. 1), it was clearly identified as the carboxy terminus of peptide 2. These results are consistent with our assignment of the first 22 amino-terminal amino acids as shown in Fig. 2.

The sequence reported here is identical to the published results for human lymphoblastoid interferon by Zoon *et al.* (4) through position 20 with two exceptions. We find threonine instead of alanine at position 14, and methionine instead of isoleucine at position 16. It is possible that these results dem-

1
Ser-Asp-Leu-Pro-Gln-Thr-His-Ser-Leu-Gly-10

11
Asn-Arg-Arg-Thr-Leu-Met-Leu-Leu-Ala-Gln-20

21
Met-Arg-

FIG. 2. Amino-terminal sequence of human leukocyte interferon. Differences between this sequence and that of lymphoblastoid interferon (up to position 20, ref. 4) are shown in boldface: Thr-14 instead of Ala-14 and Met-16 instead of Ile-16.

onstrate that leukocyte interferon differs structurally from lymphoblastoid interferon. This may result from the expression of different structural genes or possibly mutations that have become stabilized in the lymphoblastoid cells during long-term culture. The positions of nine amino acids (positions 2-6, 9-11, and 15, Fig. 2) are identical to those in mouse interferon (5).

It is noteworthy that the evidence for heterogeneity of human leukocyte interferon that we previously reported (2, 11) has been substantiated by the purification of eight distinct species of human leukocyte interferon to homogeneity (unpublished data). In general, these appear to have similar tryptic peptide maps. However, some structural differences between these species are likely. Some differences may relate to extent of glycosylation and other modifications. However, it appears that one or more internal peptides may differ in sequence (unpublished data), suggesting that two or more different alleles are expressed. Furthermore, human leukocyte interferon isolated

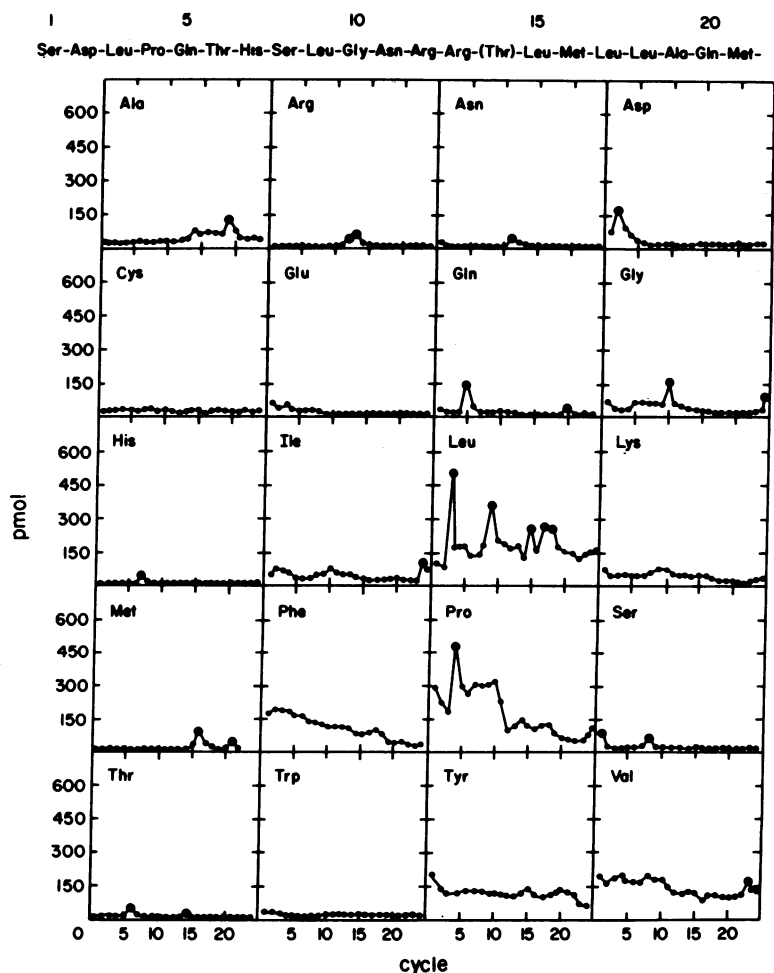


FIG. 1. Analysis of phenylthiohydantoin amino acids obtained from HPLC of each sequencer cycle of human leukocyte interferon. The tentative assignment of Thr for cycle 14 is indicated by parentheses.

from lymphoblasts (4) appears to have a structure distinct from that reported here. Accordingly, we believe that human leukocyte interferon represents a class of molecules exhibiting distinct but closely related primary amino acid sequences. Our recombinant-DNA plasmids containing leukocyte interferon sequences appear to represent distinct but homologous sequences (unpublished data). Of further note is that monoclonal antibodies to human leukocyte interferon interact differently with several of the individual purified leukocyte interferon species (T. Staehelin, B. Durrer, J. Schmidt, M. Rubinstein, W. Levy, and S. Pestka, unpublished data). All these observations lead us to postulate the existence of diversity in the primary structure of human leukocyte interferon.

In contrast, human fibroblast interferon appears to represent a single protein with a unique amino acid sequence (3, 18). Because there is likely to be some homology between them, we postulate that the structural gene for fibroblast interferon was separated from the genes for leukocyte interferon during evolution and is under different more conventional regulatory controls.

The interferon system may represent a primitive system of defense in existence prior to the evolution of the sophisticated antibody system. The diversity of leukocyte interferon may relate to optimization of its activity against different challenges (viruses or tumor cells, for example). The observations that the antiviral and antiproliferative activities of the individual species differ are consistent with the above hypothesis (19, 20). It is noteworthy that embryonic and developing cells seem to have an inactive interferon system (21); the possibility thus exists that the differential effects of the individual interferon species may play a role in differentiation and development. The exact number of human leukocyte interferon proteins that exist in specific cells and the nature of their structural and functional differences will require the elucidation of their complete primary structures and the nucleotide sequences of the corresponding cDNAs and genes.

We thank Dr. Sidney Udenfriend and Dr. Charles W. Todd for enthusiastic support of this work; Dr. Stanley Stein for making available instrumentation he developed for picomole assay of proteins, peptides, and amino acids; Dr. Robert Hershberg and Robert Bartlett for assistance in purification of interferon; Larry Brink for amino acid analysis; Louise Gerber and Dr. John Moschera for analysis of phenylthiohydantoin amino acid derivatives; and Mark J. Levy for his expert technical assistance. J.S. is a member of the City of Hope Cancer Research

Center and is supported by National Cancer Institute Grant CA16434.

1. Isaacs, A. & Lindenmann, J. (1957) *Proc. R. Soc. London Ser. B* **147**, 258-267.
2. Rubinstein, M., Rubinstein, S., Familletti, P. C., Miller, R. S., Waldman, A. A. & Pestka, S. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 640-644.
3. Knight, E., Jr., Hunkapiller, M. W., Korant, B. D., Hardy, R. W. F. & Hood, L. E. (1980) *Science* **207**, 525-526.
4. Zoon, K. C., Smith, M. E., Bridgen, P. J., Anfinsen, C. B., Hunkapiller, M. W. & Hood, L. E. (1980) *Science* **207**, 527-528.
5. Taira, H., Broeze, R. J., Jayaram, B. M., Lengyel, P., Hunkapiller, M. W. & Hood, L. E. (1980) *Science* **207**, 528-530.
6. Tarr, G. E. (1977) *Methods Enzymol.* **47**, 335-357.
7. Levy, W. P. (1981) *Methods Enzymol.*, in press.
8. Wittmann-Liebold, B. (1973) *Hoppe Seylers Z. Physiol. Chem.* **354**, 1415-1431.
9. Hunkapiller, M. W. & Hood, L. E. (1980) *Science* **207**, 523-525.
10. Rubinstein, M., Rubinstein, S., Familletti, P. C., Gross, M. S., Miller, R. S., Waldman, A. A. & Pestka, S. (1978) *Science* **202**, 1289-1290.
11. Rubinstein, M., Rubinstein, S., Familletti, P. C., Hershberg, R. D., Brink, L. D., Gutterman, J., Hester, J. & Pestka, S. (1980) in *Proteins: Structure and Biological Function*, eds. Gross, E. & Meienhofer, J. (Pierce Chemical Company, Rockford, IL), pp. 99-103.
12. Edman, P. (1950) *Acta Chem. Scand.* **4**, 283-293.
13. Zimmerman, C. L., Apella, E. & Pisano, J. J. (1977) *Anal. Biochem.* **77**, 569-573.
14. Abrahamsson, M., Gröningsson, K. & Castensson, S. (1978) *J. Chromatogr.* **154**, 313-317.
15. Bhowm, A. J., Mole, J. E., Weissinger, A. & Bennett, J. C. (1978) *J. Chromatogr.* **148**, 532-535.
16. Hunkapiller, M. W. & Hood, L. E. (1978) *Biochemistry* **17**, 2124-2133.
17. Wittmann-Liebold, B., Graffinder, H. & Kohls, H. (1976) *Anal. Biochem.* **75**, 621-633.
18. Stein, S., Kenny, C., Friesen, H.-J., Shively, J., Del Valle, U. & Pestka, S. (1980) *Proc. Natl. Acad. Sci. USA* **77**, in press.
19. Evinger, M., Rubinstein, M. & Pestka, S. (1980) in *Interferon: Properties and Clinical Uses*, ed. Khan, A. (Leland Fikes Foundation, Dallas, TX), pp. 249-263.
20. Evinger, M., Rubinstein, M. & Pestka, S. (1980) *Ann. N.Y. Acad. Sci.*, in press.
21. Burke, D. C., Graham, C. F. & Lahman, J. M. (1978) *Cell* **13**, 243-248.